

Bloch oscillations: Inverse problem

José A. González¹ · Saúl Hernández-Ortiz¹ · Carlos E. López¹ · Alfredo Raya¹

Received: date / Accepted: date

Abstract We use a neural network approach to explore the inverse problem of Bloch oscillations in a monoatomic linear chain: given a signal describing the path of oscillations of electrons as a function of time, we determine the strength of the applied field along the direction of motion or, equivalently, the lattice spacing. We find that the proposed approach has more than 80% of accuracy classifying the studied physical parameters.

Keywords Bloch Oscillations, Artificial Neural Networks, Linear chain.

1 Introduction

Understanding the motion of Bloch electrons in crystal structures is one of the traditional subjects of solid state physics, for which a number of techniques have been developed. The wave-particle duality of electrons in a crystal plays an essential role. Particularly, in the presence of an electric field, electrons are accelerated and its wavelength is shortened in such a manner that when it is twice the lattice spacing, it undergoes Bragg diffraction in reciprocal and real spaces, hence describing a periodic motion whose period depends entirely on the periodicity of the crystal along the field direction

and the strength of the applied field itself [1,2]. This periodic motion has been dubbed Bloch oscillations (BO) and its description is known for almost one century from the seminal work of Bloch [1]. In real solids, intraband tunneling and ultrafast electron scattering prevent BO from being observed. Nevertheless, in high-purity semiconductor superlattices this phenomenon has been directly observed under a variety of experimental conditions [3,4,5,6,7,8,9]. The smoking gun is a plot of a certain displacement of electron position with sinusoidal profile in time.

Moreover, within the field of quantum simulators, where the main idea is that quantum dynamics of a given system can be reproduced by an equivalent optical system which can be quantum or classical. BO can be observed in a variety of different equivalent systems to bulk crystals, like atomic systems [11,12], dielectric [13,14,15] and plasmonic [16] waveguide arrays and so on. Therefore, in this article we exploit the mathematical structure of the equations governing BO which can be extended to other systems

In this article, we discuss the *inverse* BO problem (IBO): Given an “*experimental*”, periodic curve representing the real space motion of electrons in a solid subject to an external electric field as a function of time and assuming that BO drive the dynamics, we determine the strength of the applied field or, equivalently, the lattice spacing at fixed electric field intensity. We use the linear chain as our working example and consider the nearest (1N) and next-to-nearest (2N) neighbors tight binding approximation.

Although a plethora of methods to analyze periodic curves are known and well established, in this article we construct an efficient machine learning algorithm based on Artificial Neural Networks (ANN) [17,18]. The basic idea of the ANNs consists in simulate the complex

José A. González
gonzalez@ifm.umich.mx

Saúl Hernández-Ortiz
sortiz@ifm.umich.mx

Carlos E. López
clopez@ifm.umich.mx

Alfredo Raya
raya@ifm.umich.mx

¹Instituto de Física y Matemáticas, Universidad Michoacana de San Nicolás de Hidalgo.

structure of the biological neurons with a simplified interconnection of computational nodes whose main objective is to process a set of signals and learn to classify them. When this is done in a proper way, the ANN can produce accurate predictions for signals that have never been presented to the network. This method of classification can be very efficient and accurate.

To address the issue of IBO, the remaining of the article is organized as follows: Section 2 presents the semiclassical framework to describe BO in a monoatomic lineal chain. Section 3 is devoted to introduce the setup of our ANN. Section 4 presents the data processing analysis and final remarks are given in Section 5.

2 Bloch oscillation: Semiclassical approach

We start our discussion from the tight-binding Hamiltonian of a monoatomic lineal chain of lattice spacing a . At the 1N approximation, we have

$$H\psi_n(\mathbf{k}) = -t\psi_{n+1}(\mathbf{k}) - t\psi_{n-1}(\mathbf{k}) + \epsilon_0\psi_n(\mathbf{k}) \\ \equiv \mathcal{E}^{(n)}(\mathbf{k})\psi_n(\mathbf{k}), \quad (1)$$

where t is the hopping parameter. From Bloch theorem, it is straightforward to find that in this case, the energy-momentum dispersion relation is

$$\mathcal{E}^{(n)}(\mathbf{k}) = \epsilon_0 - \epsilon^{(n)}(\mathbf{k}), \quad (2)$$

where

$$\epsilon^{(n)}(\mathbf{k}) = w(1 - \cos(|\mathbf{k}|a)), \quad (3)$$

ϵ_0 is the on-site energy and $w = 2t$. Similarly, for the 2N neighbors approximation,

$$H\psi_n(\mathbf{k}) = -t_2\psi_{n+2}(\mathbf{k}) - t_2\psi_{n-2}(\mathbf{k}) \\ - t_1\psi_{n+1}(\mathbf{k}) - t_1\psi_{n-1}(\mathbf{k}) + \epsilon_0\psi_n(\mathbf{k}) \\ \equiv \mathcal{E}^{(nn)}(\mathbf{k})\psi_n(\mathbf{k}). \quad (4)$$

In this case, t_2 and t_1 are the hopping for the corresponding neighbor. Then, from Bloch theorem we have that the dispersion relation is

$$\mathcal{E}^{(nn)}(\mathbf{k}) = \epsilon_0 - \epsilon^{(nn)}(\mathbf{k}), \quad (5)$$

but in this case

$$\epsilon^{(nn)}(\mathbf{k}) = w(1 - \cos(|\mathbf{k}|a) - w' \cos(2|\mathbf{k}|a)), \quad (6)$$

and $w = 2t_1$ while $w' = t_2/t_1$.

Next, we recall the semiclassical equations of motion for an electron moving in an external electric field \mathbf{E} oriented parallel to the linear chain,

$$\frac{d\mathbf{k}}{dt} = -e\mathbf{E}, \quad (7)$$

$$\frac{d\mathbf{r}}{dt} = \frac{1}{\hbar} \frac{\partial}{\partial \mathbf{k}} \epsilon(\mathbf{k}), \quad (8)$$

where in place of $\epsilon(\mathbf{k})$ we can insert $\epsilon^{(n)}(\mathbf{k})$ or $\epsilon^{(nn)}(\mathbf{k})$ from Eqs. (3) or (6), respectively. We can straightforwardly integrate the equations of motion and obtain the velocities and trajectories for a given external field strength. Considering the chain oriented along the x axis and a uniform electric field $\mathbf{E} = E\hat{e}_x$, we integrate Eq. (7) assuming the initial condition $k(0) = 0$. Thus

$$k(t) = -\frac{eE}{\hbar}t, \quad (9)$$

where k is the crystal-momentum of electrons in the linear chain. On the other hand, the electron velocity is given in Eq. (8), which for 1N is

$$v_{(n)}(k(t)) = \frac{wa}{\hbar} \sin(k(t)a), \\ = -\frac{wa}{\hbar} \sin\left(\frac{eEa}{\hbar}t\right). \quad (10)$$

Finally, we obtain the profile of BO simply integrating Eq. (10), which yields

$$x_{(n)}(t) = \frac{w}{eE} \cos\left(\frac{eEa}{\hbar}at\right), \\ = \frac{w}{eE} \cos(\omega_E t), \quad (11)$$

with $\omega_E = eEa/\hbar$. Analogously, for the 2N, we straightforwardly obtain

$$v_{(nn)}(t) = -\frac{wa}{\hbar} (\sin(\omega_E t) + 2w' \sin(2\omega_E t)), \quad (12)$$

$$x_{(nn)}(t) = \frac{w}{eE} (\cos(\omega_E t) + 2w' \cos^2(\omega_E t)). \quad (13)$$

Equations (11) and (13) give a fairly good agreement to the experimental observations of BO [3, 4, 5, 6, 7, 8, 9]. In the next section we introduce artificial neural networks as the tool to analyze these oscillations.

3 Artificial Neural Networks

Among the different machine learning algorithms available nowadays, the artificial neural networks algorithm is a very useful and common option to solve classification problems. In this section we describe the implementation we use to classify BO. Then we set the problem of IBO to test the performance of the approach. For simplicity, let us describe the method when we consider only the interaction of the 1N. The parameter a is fixed and the magnitude of the electric field E changes from 0.01 to 1 in steps of $\Delta E = 0.01$, i.e. $E^p = p \cdot \Delta E$ with $p = 1, \dots, P$ and $P = 100$, in units where $\hbar = 1 = e$.

We choose a discrete set of points in time t_i and a given value of the magnitude of the electric field E^p . Evaluating Eq. (11) with a and t_i given for every E^p produces a set of discrete values for the position x_i^p . The upper index p reflects the dependence of the position of

the electron on the electric field. The pairs $(E^p, x^p(t_i) = x_i^p)$ are called the training and validation sets. From the total number of generated trajectories, 70% of them are in the training set and the other 30% are in the validation set.

Now, we use an ANN to classify the BO in terms of the electric field intensity. As a first approach, we introduce the values x_i^p as inputs of the ANN and propagate them forward obtaining as a result a value that we interpret as the corresponding electric field \tilde{E}^p . A more detailed analysis shows that in order to improve the efficiency and accuracy of the network, instead of using all the values x_i^p as inputs, it is possible to compute a discrete fast Fourier transform (DFFT) of the trajectory and only use the maximum value of the amplitude and the corresponding frequency. This approach reduces the number of inputs to only two.

The structure of the ANN consists in three layers (input, hidden and output) of neurons. All the neurons of one layer are connected to neurons of the next layer with weights. Each of these weights is mathematical represented with a number, initialized randomly between -1 and 1. Every neuron has an activation function that it is used to decide how the information is propagated into the next layer. The activation functions used in this paper are logistic and linear.

After propagating forward (from the inputs to the outputs) the amplitude and frequency of the trajectory, the obtained result \tilde{E}^p in general does not correspond to the expected electric field E^p because the ANN is initialized with random coefficients. To achieve accuracy, we need to perform an optimization process where the coefficients of the network are updated in order to classify in a proper way all the elements of the training set. First, we define a cost function as

$$c = \frac{1}{2} \sum_{k=1}^K (E_k^p - \tilde{E}_k^p)^2, \quad (14)$$

where K is the number of output neurons and depends on the number of parameters predicted by the ANN.

We perform a minimization process using offline back-propagation algorithm with a learning parameter γ [17]. Given the trajectory of the BO for one value of E , the internal parameters of the ANN are adjusted such that the predicted value \tilde{E}^p gets as close as possible to E^p . Afterward, we perform the same process over the remaining elements of the training set to achieve a good accuracy predicting all the possible values of the parameter E in the sample. Next, we use a different set of trajectories, which we call validation set, to track the behavior of the training process. If the validation error starts growing or the training error reaches a given threshold, then we stop the training process and move

on to the prediction phase. After the training process is complete, we use new signals to test the accuracy of the predictions of the ANN. This set has the same amount of signals than the validation set and is created using random values for E in the given interval and it is called the prediction set. In this stage the IBO problem is posed. We are ready now to use these structures to the IBO analysis. In the next section we present technical details.

4 Data processing

The signals were discretized considering the sampling theorem [18]. This theorem establishes that the sample rate has to be at least twice the maximum frequency of the signal. In our case,

$$\omega_E = Ea, \quad (15)$$

and so we use

$$\Delta t = \frac{1}{2Ea}. \quad (16)$$

4.1 Bloch Oscillations for 1N interactions

4.1.1 Variable E (a fixed)

One hundred signals of BO were generated using Eq. (11), varying only E ($0.01 \leq E \leq 1$) with steps of $\Delta E = 0.01$. All the remaining parameters were fixed ($a = 1.0, w = 0.5$). The total number of t_i is equal to 700 to ensure that the signal created with the lowest frequency, at least has one period (Figure 1).

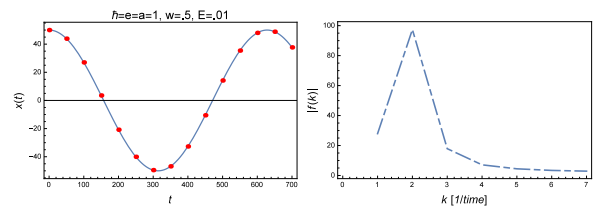


Fig. 1 Left: Trajectory for a 1N scenario corresponding to the lowest ω_E ($a = 1.0, w = 0.5, E = .01$). The dots show the positions where the DFFT is computed. Right: Absolute value of the DFFT. The selected frequency and amplitude are $k^1 = 2$ and $|f(2)|$.

A DFFT $f(k)$ is computed over each trajectory and the maximum value of the amplitude and the corresponding frequency of the signal were used as inputs for the network. It is computed as

$$f(k) = \sum_{j=0}^N x^p(t_j) \text{Exp} \left[\frac{-2\pi i k j}{N} \right], \quad (17)$$

where the indexes j and k are related with the discrete values of time and frequency. In the case of the time, we change the index i by j to avoid confusions with the complex number i appearing in Eq. (17). The relation between k and the frequency is given by $\nu_k = \frac{k}{N}2\omega_E$. We use a super-index p to denote the energy associated with the signal.

The amplitude and frequency are rescaled before using them as inputs of the ANN using the following transformation

$$\tilde{X} = \frac{X - \frac{1}{P} \sum X}{\max(X) - \min(X)}, \quad (18)$$

where the maximum and minimum values of X are computed over all the elements on the training set. Substituting k^p and $|f(k^p)|$ instead of X gives us the two required inputs for the network.

The ANN classifies the input data between ten possible ranges for the electric field. We call each of these ranges a class and they are parametrized by the index $cl = 1, \dots, 10$. This means that the output generated by the network will be associated with a class, i.e., with a corresponding electric field strength. The output neuron uses a linear activation function, so that values outside the interval $[0, 1]$ are associated with the first or last class. The value of the targets associated with each class is given by the relation

$$T_1^p = 0.1cl - 0.05, \quad (19)$$

where the numbers 0.1 and 0.05 come from the fact that we divide the output interval in 10 classes. The total number of steps used to train the network is $S = 10,000$ and the learning parameter is chosen as $\gamma = 0.0054$. The training process is repeated five times to verify a good performance and avoid strange initializations of weights. The weights obtained from the training are used to test the performance for every set.

To decide if the prediction is correct, we consider the following rule. If the electric field corresponding to the configuration lies in the range

$$T_1^p - 0.05 < E^p < T_1^p + 0.05 \quad (20)$$

then the output is taken as a correct classification, which means that the value E that generates the signal is in the range of values that corresponds to that class. Otherwise is consider a wrong classification. Once the network is completely trained over the training set, we use a set of generated trajectories with random values of E between $[0, 1]$ to verify the prediction capabilities of the ANN. The average over five different prediction sets and the efficiency of the ANN for all the sets is shown in Table 1.

	Training set	Validation set	Test set
CCP	70	30	29
Percent	1	1	0.966

Table 1 Correctly Classified Patterns (CCP) of the ANN for the signals generated varying E between $[0, 1]$. We present the results for 70 training patterns, 30 validation patterns and the average of 5 sets of prediction of 30 patterns each.

4.1.2 Variable a (E fixed).

We repeat exactly the same procedure as described above, but now fixing the value of the electric field $E = 1$ and allowing the parameter a to change between 0.505 and 1. One hundred trajectories were generated using steps of $\Delta a = 0.005$.

The classes are selected with the condition

$$\frac{cl - 1}{2} < a - 0.5 \leq \frac{cl}{20}.$$

As before, the initial weights were generated randomly between $[-1, 1]$ and the total of steps for the training was $S = 10000$ with a learning rate of $\gamma = 0.0055$. The efficiency of the ANN for this scenario is presented on Table 2.

	Training set	Validation set	Test set
CCP	65	27	28.6
Percent	0.928	0.9	0.953

Table 2 Correctly classified patterns (CCP) of the ANN for the signals generated varying a between $[0.5, 1]$. We present the results for 70 training patterns, 30 validation patterns and the average of 5 sets of prediction of 30 patterns each.

4.2 Bloch Oscillations for 2N interactions

When 2N interactions are taken into account, signals are generated using Eq. (13), varying simultaneously a and E with the remaining parameters fixed ($w = 1/2, w' = 1/4$). In these cases the network has two neurons in the output layer associated with the values of a and E . The range of the parameters studied in this scenario is $0.5333 \leq a \leq 1$ with $\Delta a = 0.0333$ and $0.0666 \leq E \leq 1$ with $\Delta E = 0.0666$. Therefore, 225 signals were generated (Figure 2) each of them with a time length of 400.

Six values are stored, three of these values corresponding to the frequencies (k_1, k_2, k_3) and the other three corresponding to their amplitudes ($|f(k_1)|, |f(k_2)|, |f(k_3)|$). As before, Eq. (18) is used to rescale the values of the amplitudes and the frequencies. Instead of using 10 classes, now we use 3 for each parameter which

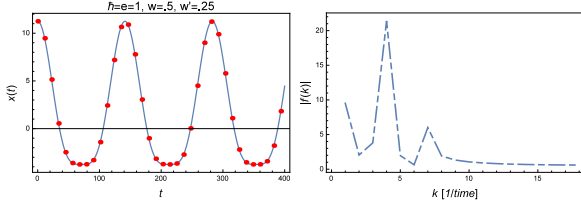


Fig. 2 *Left:* Trajectory for 2N generated using $w = 0.5$, $w' = 0.25$, $a = .66$, $E = 0.066$ corresponding to $p = 76$. The dots show the positions where the DFFT is computed. *Right:* Absolute value of the DFFT. The selected frequencies are $k_1^{76} = 1$, $k_2^{76} = 4$, $k_3^{76} = 7$ and $|f(2)|$, $|f(4)|$, $|f(7)|$ respectively.

means in total 9 classes, where the indexes are cl_a for a and cl_E for E . The number of classes was selected in such a way that the training and prediction time of the ANN is similar to the previous case. The corresponding targets are:

$$T_1^p = \frac{1}{3} \left(cl_a - \frac{1}{2} \right), \quad T_2^p = \frac{1}{3} \left(cl_E - \frac{1}{2} \right). \quad (21)$$

The parameters for the training phase are weights between $[-1, 1]$, $S = 20000$, $\gamma = .00022$ and the conditions to decide if the classification is correct are

$$T_1^p - \frac{1}{6} < O_1^p < T_1^p + \frac{1}{6}, \quad T_2^p - \frac{1}{6} < O_2^p < T_2^p + \frac{1}{6}. \quad (22)$$

The efficiency of the network for this experiment is in Table 3.

	Training set		Validation set		Test set	
Output	O_1	O_2	O_1	O_2	O_1	O_2
CCP	142	156	60	64	56.8	61.6
Percent	0.898	0.987	0.895	0.955	0.847	0.919

Table 3 Correctly classified patterns (CCP) of the ANN for the signals generated with parameters $w = 0.5$, $w' = 0.25$ and varying a between $[0.5, 1]$ and E between $[0, 1]$. We present the results for 158 training patterns, 67 validation patterns and the average of 5 training sets of 67 patterns each one.

5 Final remarks

In this article, we have carried out an analysis of BO in the linear, monoatomic chain at the nearest and next-to-nearest neighbors level within a tight-binding approach. We have developed an efficient machine learning procedure using ANN to explore the IBO problem. We develop a training process of the ANN such that, according to our error functions, the accuracy of predictions is optimized for the ANN internal parameters in a validation stage. Then the predictability of the ANN is tested directly from “experimental” curves

derived from directly integrating the equations of motion (11) and (13) for given parameters of the electric field strength E and interatomic distance a . Retaining a fixed, we have been able to predict the electric field strength of a random trajectory sample with 91% of accuracy for the 1N level. Similar accuracy of predictions was obtained for fixed E but varying interatomic distance a . When both the parameters were allowed to change, at the 2N level, our predictions are accurate up to 84% for the interatomic distance and 93% for the electric field strength. Because of the performance of the ANN, these findings are encouraging to extend a similar strategy to explore IBO in realistic systems. Graphene [19,20] and topological insulators [21] characterized by linear dispersion relations at low energies are natural candidates for this purpose. Initial work in this direction is being carried out and results will be presented elsewhere.

Acknowledgements We acknowledge support from CONA-CyT grant 256494 and CIC-UMSNH (México) under grants 4.22 and 4.23. AR acknowledges Y. Arredondo for valuable discussions.

References

1. Bloch F.: Über die Quantenmechanik der Elektronen in Kristallgittern, *Z. Phys.* **52**, 555, (1928).
2. Zener C.: Non-Adiabatic Crossing of Energy Levels, *Proc. R. Soc. A* **137**, 696, (1932).
3. Feldmann J., Leo K., Shah J., Miller D. A. B., Cunningham J. E., Meier T., von Plessen G., Schulze A., Thomas P., and Schmitt-Rink S.: Optical investigation of Bloch oscillations in a semiconductor superlattice, *Phys. Rev. B* **46** 7252 (1992).
4. von Plessen G. and Thomas P.: Method for observing Bloch oscillations in the time domain, *Phys. Rev. B* **45**, 9185 (1992).
5. Leo K., Bolivar P. H., Brüggemann F., Schwedler R., and Köhler K.: Observation of Bloch oscillations in a semiconductor superlattice, *Solid State Commun.* **84**, 943 (1992).
6. Leisching P., Haring Bolivar P., Beck. W., Dhaibi Y., Brüggemann F., Schwedler R., Kurz H., Leo K., and Köhler K.: Bloch oscillations of excitonic wave packets in semiconductor superlattices *Phys. Rev. B* **50** 14389 (1994).
7. Dekorsy T., Leisching P., Köhler K., and Kurz H.: Electro-optic detection of Bloch oscillations, *Phys. Rev. B* **50** 8106 (1994).
8. Dekorsy T., Ott R., Kurz H., and Köhler K.: Bloch oscillations at room temperature, *Phys. Rev. B* **51** 17275 (1995).
9. Waschke C., Roskos H. G., Schwedler R., Leo K., Kurz H., and Köhler K.: Coherent submillimeter-wave emission from Bloch oscillations in a semiconductor superlattice, *Phys. Rev. Lett.* **70**, 3319 (1993).
10. Roskos H. G., Waschke C., Schwedler R., Leisching P., Dhaibi Y., Kurz H., and Köhler K. Bloch oscillations in GaAs/AlGaAs superlattices after excitation well above the bandgap, *Superlattices and Microstructures* **15** 281 (1994).

11. Dahan M. B, Peik E., Reichel J, Castin Y. and Salomon, C. *Phys. Rev. Lett.***76**, 45084511 (1996).
12. Genske, M. et al., *Phys. Rev. Lett.* **110**, 190601. (2013)
13. Pertsch T, Dannberg P, Elflein W, Bräuer, A and Lederer F 1999, *Phys. Rev. Lett.* **83**.
14. Morandotti R., Peschel U., Aitchison J. S., Eisenberg H S, and Silberberg Y 1999, *Phys. Rev. Lett.* **83**, 47564759.
15. Sapienza R et al., 2003, *Phys. Rev. Lett.* **91**, 263902.
16. Block A et al, 2014, *Nat. Commun.* 53843
17. Rojas R.: *Neural Networks. A Systematic Introduction*, Springer-Verlag (1996).
18. Proakis J. G. and Manolakis D. G.: *Digital Signal Processing*, Prentice Hall (2006).
19. Novoselov K. S., Geim A. K., Morozov S. V., Jiang D., M. I. Katsnelson, Grigorieva I. V., Dubonos S. V. and Firsov A. A.: Two-dimensional gas of massless Dirac fermions in graphene, *Nature* **438**, 197 (2005).
20. Zhang Y., Tan Y. W., Stormer L. and Kim P.: Experimental observation of the quantum Hall effect and Berry's phase in graphene, *Nature* **438**, 201 (2005).
21. Vafec O. and Vishwanath A.: Dirac Fermions in Solids: From High- T_c Cuprates and Graphene to Topological Insulators and Weyl Semimetals, *Ann. Rev. Cond. Matt. Phys.* **5**, 83 (2014).